IRW

PATENT
Attorney Docket No. 0553-0118.01

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:
Shunpei Yamazaki

Serial No.: 10/694,477

Filed: October 27, 2003

Examiner: Mark V. Prenty

Art Unit: 2822

For: NONVOLATILE MEMORY AND ELECTRONIC APPARATUS

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to:

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

DATE: June 16 2006
NAME: Shannon Wallace
SIGNATURE: Shannon Wallace

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

TRANSMITTAL OF ENGLISH TRANSLATION OF PRIORITY DOCUMENTS

In furtherance of Amendment G filed May 3, 2006, in response to the Office Action of February 3, 2006, Applicant is submitting herewith a verified English translation of the priority documents: serial no. JP-09-333453 filed on November 18, 1997 in Japan; serial no. JP-09-337710 filed on November 21, 1997 in Japan; and serial no. JP-09-340754 filed on November 26, 1997 in Japan. A certified copy of each of these priority documents was already filed in the parent U.S. application 09/192,745.

Therefore, as explained in Amendment G, JP-10-093100 (a Japanese publication of the priority application for Yamazaki US 6,127,702) is not prior art to the present application.

In particular, JP '100 was published on April 10, 1998.

The present application was filed on October 27, 2003 as a divisional application under 35 USC §120 of U.S. application serial no. 09/192,745 filed on November 16, 1998 which claims the benefit under 35 USC §119 of the following foreign applications:

serial no. JP-09-333453 filed on November 18, 1997 in Japan,

serial no. JP-09-337710 filed on November 21, 1997 in Japan, and

serial no. JP-09-340754 filed on November 26, 1997 in Japan.

Therefore, as Applicants have filed a certified copy and verified English translation of these priority documents, Applicants are entitled to claim the benefit of priority date of each of these priority applications. Since each of these priority applications was filed prior to the publication date of JP '100, JP '100 is not prior art to the present application.

Accordingly, it is respectfully requested that JP-10-093100 be withdrawn as prior art to the present application.

Conclusion

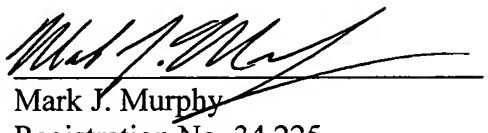
Therefore, for the reasons discussed in Amendment G, the present application is allowable over the prior art, is in a condition for allowance and should be allowed.

Applicant does not believe that an extension of time is necessary for this submission. However, if such an extension of time is needed, please consider this a petition for such an extension of time and please charge our deposit account 50/1039 for the fee for such an extension.

If any further fee is due for this submission, please charge our deposit account 50/1039.

Favorable consideration is earnestly solicited.

Respectfully submitted,



Mark J. Murphy
Registration No. 34,225

COOK, ALEX, MCFARRON, MANZO,
CUMMINGS & MEHLER, LTD.
200 West Adams Street - Ste. 2850
Chicago, IL 60606
(312) 236-8500

Customer no. 26568

Docket No.:0553-0118.01

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:)
Shunpei YAMAZAKI)
Application No.: 10/694,477) Examiner: M. PRENTY
Filed: October 27, 2003) Group Art Unit: 2822
For: NONVOLATILE MEMORY AND)
ELECTRONIC APPARATUS)

I hereby certify that this correspondence is being deposited with the United State Postal Service as first class mail in a envelope addressed to: Commissioner for Patents, P.O. Box 145C Alexandria, VA 22313-1450, on 6/16/06

Joanne Wallace

Date: June 16 2006

VERIFICATION OF TRANSLATION

Commissioner for Patents
P.O.Box 1450
Alexandria, VA 22313-1450

Sir:

I, Nami KOSAKA, C/O Semiconductor Energy Laboratory Co., Ltd. 398, Hase, Atsugi-shi, Kanagawa-ken 243-0036 Japan, a translator, herewith declare:

that I am well acquainted with both the Japanese and English Languages;

that I am the translator of the attached translation of the Japanese Patent Application No.09-333453 filed on November 18, 1997; and

that to the best of my knowledge and belief the following is a true and correct translation of the Japanese Patent Application No.09-333453 filed on November 18, 1997.

I further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Date: this 7th day of June 2006

Nami Kosaka

Name: Nami KOSAKA

[Name of Document] Patent Application

[Reference Number] P003779-01

[Filing Date] November 18, Heisei 9

[Attention] Commissioner, Patent Office

[International Patent Classification] H01L 21/00

H01L 27/10

[Title of Invention] Nonvolatile Memory and Electronic Apparatus

[Number of Claims] 6

[Inventor]

[Address] 398, Hase, Atsugi-shi, Kanagawa-ken

c/o Semiconductor Energy Laboratory Co., Ltd.

[Name] Shunpei Yamazaki

[Applicant]

[Identification Number] 000153878

[Name] Semiconductor Energy Laboratory Co., Ltd.

[Representative] Shunpei Yamazaki

[Indication of Handlings]

[Method of Payment] Prepayment

[Number of Prepayment Note] 002543

[Payment Amount] 21000

[List of Attachment]

[Attachment] Specification 1

[Attachment] Drawing 1

[Name of Document] Specification

[Title of Invention] NONVOLATILE MEMORY AND ELECTRONIC APPLIANCE

[Scope of Claims]

[Claim 1] A nonvolatile memory characterized by including a source region, a drain region, and an active region which are formed using a single crystal semiconductor; impurity regions provided locally to the active region; and an intrinsic or substantially intrinsic channel forming region interposed between the impurity regions.

[Claim 2] A nonvolatile memory characterized by including a source region, a drain region, and an active region which are formed using a single crystal semiconductor; impurity regions provided locally in the active region; and an intrinsic or substantially intrinsic channel forming region interposed between the impurity regions, wherein the impurity regions include an element selected from Group 13 or Group 15.

[Claim 3] A nonvolatile memory characterized by including a source region, a drain region, and an active region which are formed using a single crystal semiconductor; impurity regions provided locally in the active region; and an intrinsic or substantially intrinsic channel forming region interposed between the impurity regions, wherein the impurity regions includes an element selected from Group 13 or Group 15, and wherein a depletion layer expanding from the drain region toward the source region is suppressed by the impurity regions.

[Claim 4] A nonvolatile memory according to any one of claims 1 to 3, characterized in that the impurity regions are provided in a stripe form so as to reach both of the source region and the drain region.

[Claim 5] A nonvolatile memory according to any one of claims 1 to 4, characterized in that a concentration of an element included in the impurity regions ranges from 1×10^{17} to 5×10^{20} atoms/cm^3.

[Claim 6] An electronic apparatus characterized by using the nonvolatile memory according to any one of claims 1 to 5 as a storage medium.

[Detailed Description of the Invention]

[0001]

[Technical Field of the Invention]

The present invention relates to a structure of a nonvolatile memory formed by using a semiconductor. In particular, the invention is effective for a nonvolatile memory in which the channel length is 2 μm or less or moreover 0.5 μm or less.

[0002]

IC memories that perform data storing and holding in computers are generally classified into RAMs and ROMs. Examples of RAMs (random access memories) are DRAMs (dynamic RAMs) and SRAMs (static RAMs). If the power is turned off, data stored in a DRAM or an SRAM is lost.

[0003]

On the other hand, examples of ROMs (read-only memories) are mask ROMs and PROMs (programmable ROMs). Mask ROMs and PROMs have an advantage that even if the power is turned off, data stored therein is not lost. PROMs are classified into EPROMs (erasable PROMs) in which data erasure is performed by using ultraviolet light, EEPROMs (electrically EPROMs) in which data erasure is performed electrically, flash memories (flash EEPROMs) in which data erasure is performed en bloc electrically, and the like.

[0004]

To fully utilize their marked advantage of permanent data holding, researches and developments on nonvolatile memories have been made energetically. At present, the possibility of using nonvolatile memories instead of magnetic memories is being discussed.

[0005]

As for such IC memories, it is necessary to not only improve the reliability and performance but also increase the storage capacity. That is, as in the case of other types of ICs, such memory ICs are being developed according to the scaling law while miniaturization techniques are

always adopted.

[0006]

However, since basically nonvolatile memories store data according to the same principle of operation as field-effect transistors (hereinafter referred to as FETs), a short channel effect, which is known as causing serious problems in the FET operation, also causes serious problems in the operation of nonvolatile memories as the miniaturization advances.

[0007]

In particular, a phenomenon called punch-through decreases the source-drain withstanding voltage and hence makes the current control with the gate electrode difficult. A SSW-DSA structure (Nikkei Microdevices, pp. 47-48, May issue, 1992) is a conventional example of increasing the punch-through resistance.

[0008]

[Problems to Be Solved by the Invention]

In the field of the FET, the SSW-DSA structure is a structure that utilizes a technique called a pocket structure in which an impurity region having the same conductivity type as the substrate is provided in a channel-drain junction portion. This structure can prevent the occurrence of a punch-through phenomenon by suppressing the expansion of a drain depletion layer.

[0009]

However, in nonvolatile memories, electron-hole pairs are generated by positively causing impact ionization in the channel-drain junction portion. Therefore, a large amount of holes flow to the substrate side at the same time as the injection of electrons into a floating gate.

[0010]

However, in the SSW-DSA structure, a large amount of holes thus generated act in no other way than flow into a substrate terminal. This may cause a problem that a parasitic bipolar is formed among source-substrate-drain to cause a kink phenomenon (an abnormal increase in drain current).

[0011]

The present invention has been made in view of the above problems, and an object of the invention is to realize a high-performance memory by effectively preventing or suppressing the short channel effect that occurs in miniaturizing nonvolatile memories.

[0012]

[Means for Solving the Problems]

A structure of the invention disclosed in this specification is characterized by including a source region, a drain region, and an active region that are formed by using a single crystal semiconductor; stripe-shaped impurity regions provided in the active region; and an intrinsic or substantially intrinsic channel forming region interposed between the impurity regions.

[0013]

Another structure of the invention disclosed in this specification is characterized by including a source region, a drain region, and an active region that are formed by using a single crystal semiconductor; stripe-shaped impurity regions provided in the active region; and an intrinsic or substantially intrinsic channel forming region interposed between the impurity regions, wherein the impurity regions include an element selected from Group 13 or Group 15.

[0014]

Another structure of the invention disclosed in this specification is characterized by including a source region, a drain region, and an active region that are formed by using a single crystal semiconductor; stripe-shaped impurity regions provided in the active region; and an intrinsic or substantially intrinsic channel forming region interposed between the impurity regions, wherein the impurity regions include an element selected from Group 13 or Group 15, and wherein the impurity regions suppress a depletion layer expanding from the drain region toward the source region.

[0015]

In the above structure, the impurity regions are preferably provided so as to reach both of the

source region and the drain region.

[0016]

In the above structure, the concentration of the element included in the impurity regions preferably ranges from 1×10^{17} to 5×10^{20} atoms/cm^3.

[0017]

Moreover, it is effective that a storage circuit in which a nonvolatile memory including the above structure is used as a storage medium is formed and it is incorporated into an electronic apparatus.

[0018]

The main feature of the invention is that impurity regions are formed locally in the active region and the impurity regions prevent a depletion layer from expanding from the drain region toward the source region. In this specification, a region that is enclosed by a source region, a drain region, and field oxide films is called an active region and the active region is divided into stripe-shaped impurity regions and channel forming regions.

[0019]

Since the effect of preventing expansion of a depletion layer looks like pinning the depletion layer, the inventors define the term "pinning" as meaning "prevention".

[0020]

[Embodiment Mode of the Invention]

An embodiment mode of the present invention will be described in detail by using embodiments hereinafter shown.

[0021]

[Embodiments]

[Embodiment 1]

This embodiment will be described with reference to Fig. 1, which is a top view and sectional views of a non-volatile memory to which the invention is applied. This embodiment is

described taking up an EEPROM having a basic stacked structure as an example.

[0022]

In Fig. 1, reference numeral 101 denotes single crystal silicon (p-type silicon); 102, field oxide films formed by a LOCOS method; 103, a source region to which arsenic (or phosphorus) is added; and 104, a drain region. Although this is a structural example of an n-type EEPROM, it is also possible to construct a p-type EEPROM. A p-type EEPROM can be constructed by forming source and drain regions by adding boron to an n-type silicon.

[0023]

Reference numeral 105 denotes impurity regions (hereinafter referred to as pinning regions) that are the most important feature of the invention. The pinning regions 105 are formed by adding an impurity having the same conductivity type as the silicon substrate 101. In the case of Fig. 1, since the p-type silicon is used, the pinning regions 105 are formed by adding an element selected from Group 13 (typically boron). Naturally, when an n-type silicon is used (a p-type EEPROM is to be manufactured), pinning regions may be formed by adding an element selected from Group 15.

[0024]

The element selected from Group 13 or Group 15 shifts the energy band of single crystal silicon, thereby forming an energy barrier to carriers (electrons or holes). In this sense, the pinning regions can also be called energy-band-shifted regions. Elements other than Group 13 and Group 15 can also be used as long as they have such an effect.

[0025]

An energy-band-shifting element will be described with reference to schematic diagrams of Fig. 2. Fig. 2(A) shows an energy band state of single crystal silicon. If an impurity element (an element selected from Group 13) that shifts the energy band in such a direction as to obstruct the movement of electrons is added thereto, it is changed to an energy state as shown in Fig. 2(B).

[0026]

In the added region, although no change occurs in the energy band gap, the Fermi level (Ef) is moved to the valence band (Ev) side. As a result, the energy state is shifted upward in appearance and hence an energy barrier which is higher than the undoped regions by ΔE (for electrons) is formed.

[0027]

If an impurity element (an element selected from Group 15) that shifts the energy band in such a direction as to obstruct the movement of holes is added, it is changed to an energy state as shown in Fig. 2(C).

[0028]

In this case, in the added region, the Fermi level is moved to the conduction band (Ec) side. As a result, the energy state is shifted downward in appearance and hence an energy barrier which is higher than the undoped regions by ΔE (for holes) is formed.

[0029]

As described above, an energy difference corresponding to ΔE is generated between the (undoped) regions where the impurity is not added and the pinning regions 105. The height of the energy (i.e., potential) barrier depends on the concentration of the added impurity element. In the invention, the impurity element concentration is adjusted in a range of 1×10^{17} to 5×10^{20} atoms/cm^3 (preferably 1×10^{18} to 5×10^{19} atoms/cm^3).

[0030]

Since the formation of the pinning regions 105 is enabled by microprocessing technology, it is necessary to use an adding method suitable for microprocessing, such as an ion implantation method or an FIB (Focusd Ion Beam). When an adding method using a mask is employed, it is desirable to use microprocessing, for instance, forming a mask pattern by electron beam lithography.

[0031]

Most typically, the pinning regions 105 are formed in such a manner that the pinning regions 105 are approximately parallel with channel forming regions 106 and the pinning regions 105 and the channel forming regions 106 are arranged alternately as shown in Fig. 1(A). That is, it is preferable to form a plurality of stripe-shaped pinning regions 105 in a region (active region) enclosed by the source region 103, the drain region 104, and the field oxide films 102.

[0032]

It is effective to form pinning regions at side end portions of the active region (end portions where the active region is in contact with the field oxide films 102). The formation of the pinning regions at the side end portions can reduce leak current that passes through the side end portions.

[0033]

It is sufficient to form the pinning regions 105 so that they at least reach a junction portion of the active region and the drain region 104 (i.e., a drain junction portion). The effect can be obtained by preventing the expansion of the depletion layer at the drain junction portion because the depletion layer that causes the punch-through phenomenon develops from the drain junction portion. That is, the expansion of the depletion layer can be suppressed by forming dot-shaped or elliptical pinning regions 105 in the active region so that a part of them exists in the drain junction portion.

[0034]

Naturally, a pinning effect can be obtained more effectively by forming the pinning regions so that they reach both of the source region 103 and the drain region 104 as shown in Fig. 1(A).

[0035]

It is desirable that the implantation depth of the pinning regions 105 be deeper than the junction depth of the source and the drain regions. Therefore, it is necessary the implantation depth be 0.1 to 0.5 μm (preferably 0.2 to 0.3 μm).

[0036]

Now, the channel length and the channel width will be defined with reference to Fig. 3. In Fig. 3, the distance between a source region 301 and a drain region 302 (i.e., the length of an active region 303) is defined as a channel length (L). The invention is effective in a case where this length is 2 μm or less, typically 0.05 to 0.5 μm, and preferably 0.1 to 0.3 μm. The direction along this channel length is called a channel length direction.

[0037]

The width of an arbitrary pinning region 304 is called a pinning width (v_j). The pinning width may be set to 1 μm or less, typically 0.01 to 0.2 μm and preferably 0.05 to 0.1 μm. An effective pinning width (V) that is the sum of the widths of all pinning regions existing in the active layer 303 is defined as the following formula.

[0038]

[Formula 1]

[0039]

To obtain the pinning effect, it is necessary to form at least one pinning region in the active region 303; that is, a condition j = 1 or more needs to be satisfied. When pinning regions are formed at the side end portions (portions in contact with field oxide films) of the active region 303, at least a condition j = 2 or more needs to be satisfied.

[0040]

The width of a channel forming region 305 is called a channel width (w_i). The channel width can be set in accordance with any case. For memories in which large current is not required, the channel width may be set to 1 μm or less, typically 0.05 to 0.5 μm, and preferably 0.1 to 0.3 μm.

[0041]

An effective channel width (W) that is the sum of all channel widths (w_i) is defined as the following formula.

[0042]

[Formula 2]

[0043]

When the pinning regions are formed only at the side end portions of the active region 303, i = 1 is satisfied. A pinning effect can be obtained more effectively by forming pinning regions also in regions other than the side end portions of the active region 303. In such a case, i becomes 2 or more.

[0044]

A total channel width (W_{total}) that is the sum of the sum of the pinning regions (i.e., the effective pinning width) and the sum of the channel forming regions (i.e., the effective channel width) is defined as the following formula.

[0045]

[Formula 3]

[0046]

The total channel width (W_{total}) corresponds to the width of the active region 303 (i.e., the length of the active region in the direction perpendicular to the channel length direction). The direction along the total channel width is called a channel width direction.

[0047]

Since the invention is intended to be applied to nonvolatile memories having an extremely short channel length as mentioned above, the pinning regions and the channel forming regions need to be formed so as to have extremely small dimensions.

[0048]

Returning to Fig. 1, it is preferable that the impurity element that has been added to the pinning region 105 be activated by furnace annealing, laser annealing, lamp annealing, or the like. The activation step may be performed at the same time as an annealing process in a later step such as a step of forming a gate insulating film, or independently of such annealing.

[0049]

The invention is characterized in that pinning regions are formed locally (i.e., into a stripe shape) in a region of a conventional nonvolatile memory that functions as a channel forming region. Therefore, the other structures of the conventional nonvolatile memory can be employed as they are.

[0050]

That is, a tunnel oxide film 107 is also formed on the single crystal silicon where the source region 103, the drain region 104, and the pinning regions 105 are formed. The tunnel oxide film 107, which is formed by a thermal oxidation step, is required to have high film quality. In this embodiment, the thickness of the tunnel oxide film 107 is set to 11 nm. It goes without saying that the thickness of the tunnel oxide film is not limited to this value.

[0051]

In this embodiment, the pinning regions 105 may be formed even after the formation of the tunnel oxide film 107.

[0052]

A floating gate 108 including a first polysilicon layer, a first interlayer film 109, a control gate 110 including a second polysilicon layer, a second interlayer film 111, and a bit line 112 are provided on the tunnel oxide film 107.

[0053]

Of course, it is also possible to use a conductive film such as a metal film instead of the polysilicon layer. Moreover, it is effective to use a multilayer film expressed by $SiO_2/SiN/SiO_2$ (generally called an ONO film) as the interlayer film.

[0054]

The two-layer polysilicon EEPROM of this embodiment is expressed as a circuit diagram shown in Fig. 1(D), in which Vd denotes a drain voltage, Vs denotes a source voltage, C.G. denotes a control gate voltage, and F.G. denotes a potential of the floating gate.

[0055]

In the EEPROM of this embodiment, the following voltages are applied at the time of data writing and erasure.

[0056]

[Table 1]

[0057]

Naturally, the operation voltages are not limited to the values of Table 1. Further, the invention is not limited to the structure of this embodiment and can be applied to any kinds of EEPROMs in which data is erased electrically.

[0058]

[Operation Effect of the Invention]

A first effect of the invention will be described. In Fig. 1, the pinning regions 105 that are formed locally in the active region serve as stoppers with respect to the depletion layer that expands from the drain side and effectively suppress the expansion of the depletion layer. Therefore, the punch-through phenomenon that is caused by expansion of the depletion layer can be prevented. Further, since an increase of depletion layer charge due to the expansion of the depletion layer is suppressed, a reduction in threshold voltage can be prevented.

[0059]

Next, a second effect will be described. In the invention, the narrow channel effect can be enhanced intentionally by the pinning regions. The narrow channel effect, which is a phenomenon observed when the channel width is extremely narrow, causes an increase in threshold voltage (refer to Koyanagi Mitsumasa et al., Submicron Devices I, pp. 88-138, Maruzen Co., Ltd., 1987).

[0060]

Fig. 4 shows an energy state (potential state) of the active region when the pinning TFT of this embodiment operates. In Fig. 4, regions 401 and 402 correspond to the energy state of the

pinning regions 105 and a region denoted by 403 corresponds to the energy state of a channel forming region 106.

[0061]

As seen from Fig. 4, the pinning regions 105 form high-energy barriers and the channel forming regions 106 form regions with low-energy barriers. Therefore, carriers move through the channel forming regions 106 with priority where the energy state is low.

[0062]

In this manner, high-energy barriers are formed in the pinning regions 105 and the threshold voltage increases in that portion. As a result, a threshold voltage that is observed as a whole also increases. This narrow channel effect becomes more remarkable as the effective channel width decreases.

[0063]

As described above, in the invention, since the concentration of an impurity that is added to the pinning regions 105 and the effective channel width W can be designed freely, the threshold voltage can be adjusted by controlling the degree of the narrow channel effect. That is, by controlling the pinning effect, the threshold voltage can be adjusted to a desired value by balancing a threshold voltage decrease due to the short channel effect with a threshold voltage increase due to the narrow channel effect.

[0064]

Since a Group-13 element is added to the pinning regions in the case of an n-type and a Group-15 element is added in the case of a p-type, the threshold voltage is shifted in that portion in a direction in which its threshold voltage increases (in the positive direction in the case of an n-type and in the negative direction in the case of a p-type). Since the threshold voltage increases locally, the threshold voltage as a whole increases accordingly. Therefore, to adjust the threshold voltage to a desired value, it is important to set the concentration of an impurity that is added to the pinning regions at a proper value.

[0065]

Incidentally, in nonvolatile memories, discrimination between "0" and "1" is made by changing the threshold voltage by injecting charge (mainly electrons) into the floating gate and detecting whether current flows through the bit line when a certain voltage is applied. Therefore, if the threshold voltage is made unduly small by the short channel effect, it becomes necessary to discriminate between "0" and "1" by applying a very small voltage. In this case, the device becomes prone to be influenced by noise or the like and the possibility of occurrence of an erroneous operation increases.

[0066]

In contrast, in the invention, since the threshold voltage can be controlled to have a desired value by suppressing a threshold voltage reduction, the capability of discriminating between "0" and "1" is increased. That is, the invention can realize a nonvolatile memory having very high reliability.

[0067]

Next, a third effect will be described. The nonvolatile memory according to the invention has an advantage that majority carriers (electrons in the case of an n-type and holes in the case of a p-type) move through the channel forming regions 106 that are composed of substantially intrinsic regions.

[0068]

The term "substantially intrinsic region" basically means an undoped single crystal semiconductor region, and includes a region where conductivity type compensation is made intentionally by adding an impurity element of the opposite conductivity type, and a region having one conductivity type in a range where the threshold voltage can be controlled.

[0069]

For example, it can be said that a silicon wafer having a dopant concentration of 5×10^{16} atoms/cm^3 or less (preferably 5×10^{15} atoms/cm^3 or less) and carbon, nitrogen, and oxygen

concentrations of 2×10^{18} atoms/cm^3 or less (preferably 5×10^{17} atoms/cm^3 or less) is substantially intrinsic. In this sense, silicon wafers commonly used are substantially intrinsic unless an impurity is added intentionally in a process.

[0070]

When a carrier movement region is substantially intrinsic, a mobility reduction due to the impurity scattering is extremely small and hence high carrier mobility is obtained. Therefore, the carrier movement is dominated by the influence of the lattice scattering and a state that is very close to the ideal state is obtained.

[0071]

When the linear pinning regions are formed so as to reach both of the source region and the drain region as shown in Fig. 1(A), an effect is obtained that movement paths of majority carriers are defined by the pinning regions.

[0072]

As described above, each channel forming region interposed between pinning regions has an energy state as shown in Fig. 4. The structure of Fig. 1A is regarded as an arrangement of a plurality of slits each having the energy state of Fig. 4.

[0073]

Fig. 5 schematically illustrates such a state. In Fig. 5, reference numerals 501 and 502 denote pinning regions and channel forming regions, respectively. As shown in Fig. 5, majority carriers (electrons or holes) 503 cannot go over the pinning regions 501 and hence move through the channel forming regions 502 with priority. That is, the movement paths of majority carriers are defined by the pinning regions.

[0074]

Defining the movement paths of majority carriers decreases scattering due to self-collision of carriers, which greatly contributes to mobility increase. Further, since only a very small amount of impurity elements exist in the substantially intrinsic channel forming regions, the velocity

overshoot effect occurs, which is a phenomenon that the electron mobility becomes higher than usual even at room temperature (refer to K. Ohuchi et al., Jpn. J. Appl. Phys. 35, pp. 960, 1996). Therefore, the mobility becomes extremely high.

[0075]

High carrier mobility that is obtained as described above is effective in shortening the charge write time and the charge read time, thereby increasing performance of the memory function. High carrier mobility means to have high energy, and hence the charge writing efficiency is greatly increased by channel hot electron injection (CHE injection).

[0076]

Next, a fourth effect of the invention will be described. When the structure of the invention is employed, an electric field is concentrated to a large extent at the junction portions (typically, a p^+/n^{++} junction or an n^+/p^{++} is formed) of the pinning regions and the drain region. Therefore, there occur a large amount of electrons that are given high energy through acceleration or generated by impact ionization (collectively called hot electrons).

[0077]

Therefore, charge injection into the floating gate is performed very efficiently and hence the data write time can be shortened. In this way, by providing the pinning regions, the efficiency of hot electron injection at the drain junction portion can be increased.

[0078]

Next, a fifth effect will be described. The fact that the pinning regions of the invention have the functions of preventing the short channel effect and controlling the threshold voltage has been described above. In addition, the pinning regions of the invention have a very important role in preventing a parasitic bipolar from being made conductive due to impact ionization (collisional ionization).

[0079]

Conventionally, electrons of electron-hole pairs generated by impact ionization are injected

into the floating gate. On the other hand, holes flow into the substrate and cause a substrate current, which makes a parasitic bipolar conductive.

[0080]

In contrast, in the invention, holes generated by impact ionization immediately move into the pinning regions and extracted to the source region via the internal portion. Therefore, the holes do not make a parasitic bipolar conductive and hence do not lower the source-drain withstanding voltage.

[0081]

It goes without saying that this effect is particularly remarkable when the pinning regions are formed so as to reach both of the source and drain regions. Holes can be extracted more efficiently if the pinning regions are in contact with a pickup electrode in the source region.

[0082]

[Embodiment 2]

The EEPROM of a two-layer polysilicon type shown in Embodiment 1 can be discriminated as a bite erasing type (data erasure is carried out for each unit memory element) and a flash type (data erasure is performed en bloc for a block of memory elements).

[0083]

A flash type EEPROM is also called a flash memory, and the invention can be applied to either type of EEPROMs.

[0084]

Moreover, a method of data erasure includes a source erasure type, a source/gate erasure type, a substrate erasure type, or various methods. The invention can be applied to either case.

[0085]

[Embodiment 3]

While Embodiments 1 and 2 were directed to the two-layer polysilicon EEPROM, this embodiment is directed to a case where the invention is applied to a three-layer polysilicon

EEPORM. This embodiment will be described with reference to Fig. 6.

[0086]

Since an EEPROM of this embodiment has the same basic structure as the two-layer polysilicon EEPROM described in Embodiment 1, the reference numerals used in Fig. 1 are also used. That is, for the parts shown in Fig. 6 that are given the same reference numerals as in Fig. 1, reference is made to the descriptions that were made above in connection with Fig. 1. In this embodiment, only the parts that are different from that in the Embodiment 1 will be given new reference numerals and described below.

[0087]

Fig. 6(A) is different from Fig. 1(A) in that an erasing gate 601 is provided. That is, a first polysilicon layer constitutes the erasing gate 601 and second and third polysilicon layers constitute the floating gate 108 and the control gate 110, respectively.

[0088]

In the EEPROM with the structure according to Embodiment 1, data erasure is performed by extracting, to the substrate side (the source region or the bulk substrate), electrons that have been injected into the floating gate 108. In contrast, in the structure according to this embodiment, data erasure is performed by extracting, to the erasing gate 601, electrons that have been injected into the floating gate 108.

[0089]

Therefore, in Fig. 1(B), an insulating film 602 for insulating the erasing gate 601 and the floating gate 108 from each other should be as thin as possible (preferably 8 to 12 nm) so as to allow a flow of tunnel current (Fowler-Nordheim current) as well as should be of such high quality as to be highly durable.

[0090]

The EEPROM of this embodiment can be manufactured basically by the same process as that of Embodiment 1 with an exception that a step of forming the erasing gate 601 and the erasing

gate insulating film 602 after formation of the pinning regions is added.

[0091]

An EEPROM having an erasing gate like the one in this embodiment is shown as a circuit diagram in Fig. 6(D), in which Vd denotes a drain voltage, Vs denotes a source voltage, E.G. denotes an erasure gate voltage, C.G. denotes a control gate voltage, and F.G. denotes a floating gate potential.

[0092]

In the EEPROM of this embodiment, the following voltages are applied at the time of data writing and erasure.

[0093]

[Table 2]

[0094]

Naturally, the operation voltages are not limited to the values of Table 2. Further, the invention is not limited to the structure of this embodiment and can be applied to any kinds of EEPROMs having an erasing gate structure.

[0095]

[Embodiment 4]

The nonvolatile memories according to Embodiments 1 to 3 utilize hot electron injection and Fowler-Nordheim current for data writing and erasure, respectively. In contrast, Fowler-Nordheim current may be used for data writing.

[0096]

In particular, in manufacturing a large-capacity memory of 256 Mbits or more, to improve the reliability (i.e., to elongate the life by suppressing deterioration), it is preferable to write data by using Fowler-Nordheim current.

[0097]

[Embodiment 5]

Embodiment 1 was directed to the two-layer polysilicon EEPROM in which data erasure is performed electrically. On the other hand, the nonvolatile memory in which electrons that have been injected into the floating gate are extracted to the source or the substrate by exciting those by ultraviolet illumination or heating is called an EPROM. The invention can also be applied to such an EPROM.

[0098]

Among various kinds of EPROMs is a nonvolatile memory not using a floating gate in which a two-layer gate insulating film is provided between a control gate and a silicon substrate and hot electrons are captured by interface states of the two-layer gate insulating film. For example, a type in which hot carriers are captured at the interface between a silicon oxide film and a silicon nitride film is called an NMOS nonvolatile memory.

[0099]

Further, there are nonvolatile memories in which hot carriers are captured by metal clusters, silicon clusters, or the like that are formed intentionally at an interface between insulating films.

[0100]

The invention can be applied to all types of EPROMs as exemplified above.

[0101]

[Embodiment 6]

Since the invention is applicable to all kinds of conventional nonvolatile memories, it can be applied to all the known circuit configurations. This embodiment is directed to a case where the invention is applied to flash memories that are designed according to the NAND and NOR type architectures, respectively.

[0102]

First, description will be made of the configuration of a NAND-type memory circuit shown in Figs. 7(A) and (B). Although Fig. 7 shows two circuits each including eight memory transistors and two selection transistors, only one of them will be described.

[0103]

As shown in Fig. 7(A), 701 and 702 are selection transistors, which respectively have selection lines S1 and S2 shown as 703 and 704 as gate electrodes. The selection transistor 701 connects a bit line 705 shown as B1 (or B2) to eight memory transistors 706 to 713.

[0104]

Although this embodiment is directed to a case where eight memory transistors are connected to each other in series, the number of memory transistors is not limited to 8.

[0105]

The selection transistor 702 is connected to a final-stage memory transistor 713. The other terminal of the selection transistor 702 is grounded. Even if it is connected to a power supply line instead of the ground, the memory circuit can still operate properly.

[0106]

The memory transistors 706 to 713 use word lines 714 to 721 (shown as W1 to W8) as control gates, respectively.

[0107]

Fig. 7(B) is a schematic circuit pattern of the NAND-type memory circuit of Fig. 7(A). In the memory transistors, regions shown with hatched lines are floating gates that are provided under the respective control gates 714 to 721.

[0108]

Next, description will be made of the configuration of a NOR-type memory circuit shown in Figs. 8(A) and (B). Although two circuits each including four memory transistors are described in Fig. 8, description is made of only one of them.

[0109]

As shown in Fig. 8(A), four memory transistors 802 to 805 are individually connected to a bit line 801 shown as B1. The terminals (source regions) of the respective memory transistors 802 to 805 which are not connected to the bit line 801 are connected to a ground line 806.

[0110]

The memory transistors 802 to 805 use word lines 807 to 810, which are shown as W1 to W4, as control gates, respectively.

[0111]

Fig. 8(B) is a schematic circuit pattern of the NOR-type memory circuit of Fig. 8(A). In the memory transistors, regions shown with hatched lines are floating gates that are provided under the control gates 807 to 810, respectively.

[0112]

Although NAND-type circuits as shown in Fig. 7 have disadvantages that the order of writing is fixed and the read access time is long, they have an advantage that the integration density can greatly be increased.

[0113]

The configuration of the NOR-type circuit shown in Fig. 8 is effective in injecting electrons precisely into the floating gates and reading out charge amounts precisely. This is the feature of the NOR architecture in which individual memory transistors are directly connected to a source-drain bus line.

[0114]

Although this embodiment is directed to the circuits that use the nonvolatile memory having an electrode with the two-layer structure (polysilicon or the like), it is possible to carry out even in a nonvolatile memory having an electrode with the three-layer structure (the structure including the erasing gate) as described above in Embodiment 3.

[0115]

[Embodiment 7]

This embodiment is directed to a case where a nonvolatile memory according to the invention is applied to a microprocessor that is integrated on one chip, such as a RISC processor or an ASIC processor.

[0116]

Fig. 9 shows an example of a microprocessor, which is typically composed of a CPU core 11, a flash memory 12 (or a RAM), a clock controller 13, a cache memory 14, a cache controller 15, a serial interface 16, an I/O port 17, and the like.

[0117]

Of course, the microprocessor of Fig. 9 is a simplified example. Naturally, a variety of circuit designs are employed in actual microprocessors in accordance with their uses.

[0118]

In the microprocessor of Fig. 9, the CPU core 11, the clock controller 13, the cache controller 15, the serial interface 16, and the I/O port 17 are constituted of CMOS circuits 18. Pinning regions 19 according to the invention are formed in the CMOS circuits 18.

[0119]

In this manner, the invention can be applied to MOSFETs as well as nonvolatile memories, as disclosed in detail in Japanese Patent Application No. Hei. 8-232553.

[0120]

The flash memory 14 uses a nonvolatile memory according to the invention that forms a memory circuit 20. Every memory cell that constitutes the memory circuit 20 is provided with pinning regions 21. It is possible to use a nonvolatile memory according to the invention also in the cache memory 12.

[0121]

As described above, in the example of Fig. 9, the pinning technology disclosed in the invention is utilized in all of the memory portions and the logic portions.

[0122]

As occasion demands, a configuration shown in Fig. 10 may be employed. Fig. 10 shows a case where the logic portions (excluding the memory portions) are formed by ordinary CMOS circuits 22. This configuration may be implemented by not forming pinning regions only in the

logic portions.

[0123]

In this manner, it is possible to provide, at the circuit designing stage, pinning regions at portions where the pinning regions should be formed; a party who practices the invention may determine at will whether to form pinning regions in the entire circuit or only part of it. When the invention is applied to a hybrid IC in which various functions are combined, such a high degree of freedom in circuit design is very effective.

[0124]

[Embodiment 8]

A semiconductor circuit (memory circuit) formed by a nonvolatile memory according to the invention can be incorporated, as a recording medium for data storage and readout, in electronic apparatuses of every field. In this embodiment, examples of those electronic apparatuses will be described with reference to Fig. 11.

[0125]

Electronic apparatuses which can use a nonvolatile memory of the invention are a video camera, an electronic still camera, a projector, a head-mounted display, a car navigation, a personal computer, portable information terminals (a mobile computer, a cellular telephone, a PHS, etc.) and the like.

[0126]

Fig. 11(A) shows a cellular telephone, which is composed of a main body 2001, a voice output portion 2002, a voice input portion 2003, a display device 2004, manipulation switches 2005, and an antenna 2006. The invention is incorporated in an LSI substrate and used to add an address function for recording telephone numbers.

[0127]

Fig. 11(B) shows a video camera, which is composed of a main body 2101, a display device 2102, a sound input portion 2103, manipulation switches 2104, a battery 2105, and an image

receiving portion 2106. The invention is incorporated in a built-in LSI substrate and used for such a function as storage of image data.

[0128]

Fig. 11(C) shows a mobile computer (mobile computer), which is composed of a main body 2201, a camera portion 2202, an image receiving portion 2203, a manipulation switch 2204, and a display device 2205. The invention is incorporated in a built-in LSI substrate and used for storage of processed data and image data.

[0129]

Fig. 11(D) shows a head-mounted display, which is composed of a main body 2301, display devices 2302, and a band portion 2303. The invention is used in an image signal correction circuit that is connected to the display devices 2302.

[0130]

Fig. 11(E) shows a rear type projector, which is composed of a main body 2401, a light source 2402, a display device 2403, a polarizing beam splitter 2404, reflectors 2405 and 2406, and a screen 2407. The invention can be used as a storage circuit for storing data to be supplied to a γ-correction circuit.

[0131]

Fig. 11(F) shows a front type projector, which is composed of a main body 2501, a light source 2502, a display device 2503, an optical system 2504, and a screen 2505. The invention is used as a storage circuit for storing data to be supplied to a γ -correction circuit.

[0132]

As described above, the application range of the invention is extremely wide and the invention can be applied to electronic apparatuses of every field. In addition to the above examples, a nonvolatile memory of the invention can be used as a storage medium that is indispensable in various control circuits and information processing circuits.

[0133]

[Effect of the Invention]

The invention makes it possible to minimize influences of the miniaturization effects as typified by the short channel effect and the like and to further advance the miniaturization of nonvolatile memories.

[0134]

The invention also makes it possible to realize small-area and large-capacity nonvolatile memories while securing their high reliability.

[Brief Description of the Drawings]

[Fig. 1] drawings for showing the structure of a nonvolatile memory according to the present invention.

[Fig. 2] drawing for showing changes of an energy band.

[Fig. 3] a drawing for describing definitions of a channel length and a channel width.

[Fig. 4] a drawing for showing an energy state of an active region.

[Fig. 5] a drawing for showing an energy state of an active region.

[Fig. 6] drawings for showing the structure of a nonvolatile memory according to the present invention.

[Fig. 7] drawings for showing a circuit using a nonvolatile memory according to the present invention.

[Fig. 8] drawings for showing a circuit using a nonvolatile memory according to the present invention.

[Fig. 9] a drawing for showing a semiconductor circuit using a nonvolatile memory according to the present invention.

[Fig. 10] a drawing for showing a semiconductor circuit using a nonvolatile memory according to the present invention.

[Fig. 11] drawings for showing electronic apparatuses each using a nonvolatile memory

according to the present invention.

[Name of Document] Abstract

[Abstract]

[Object] To effectively suppress a short channel effect caused with miniaturization of a nonvolatile memory and achieve a high-performance memory.

[Solving Means] In a nonvolatile memory, a pinning region 105 is locally provided in an active region that is enclosed by a field oxide film 102, a source region 103, and a drain region 104. According to the present invention, a depletion layer expanding from a drain side toward a source side is suppressed by the pinning region 105 and a punch-through phenomenon due to the short channel effect is prevented.

[Selected Figure] Fig. 1

Reference No. P003779-01
Document where chemical formulas etc. are descibed
Name of Document Specification

[Formula 1]

$$V = \sum_{j=1}^{n} v_j$$

[Formula 2]

$$W = \sum_{i=1}^{m} w_i$$

[Formula 3]

$$W_{total} = V + W$$

Reference No. P003779-01
Document where chemical formulas etc. are descibed
Name of Document Specification

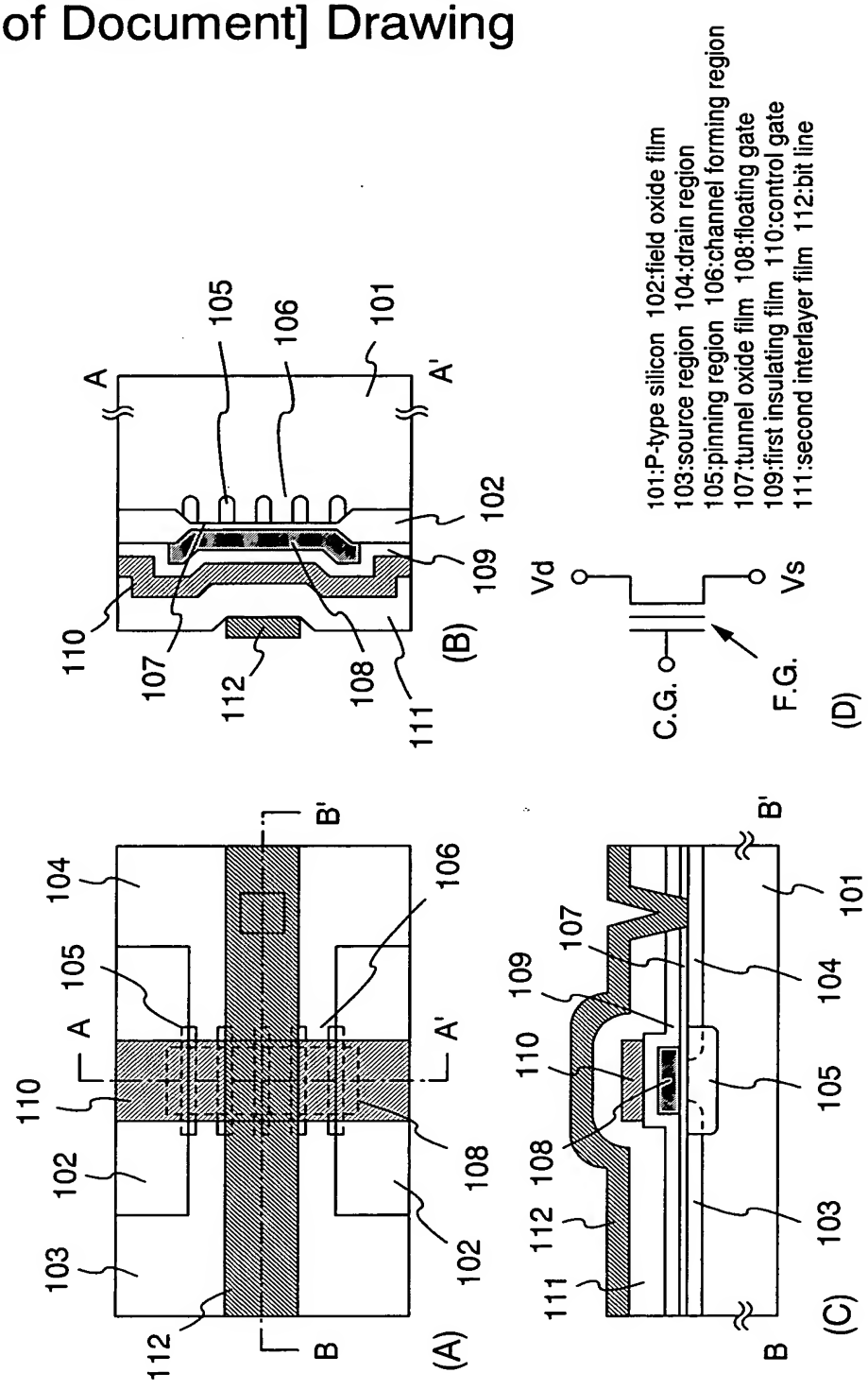
[Table 1]

mode	Vd	V_{CG}	Vs	mechanism
at writing	6	12	0	hot electron injection
at erasing	-	0	12	F-N tunnel erasure
at readout	~1	5	0	-

[Table 2]

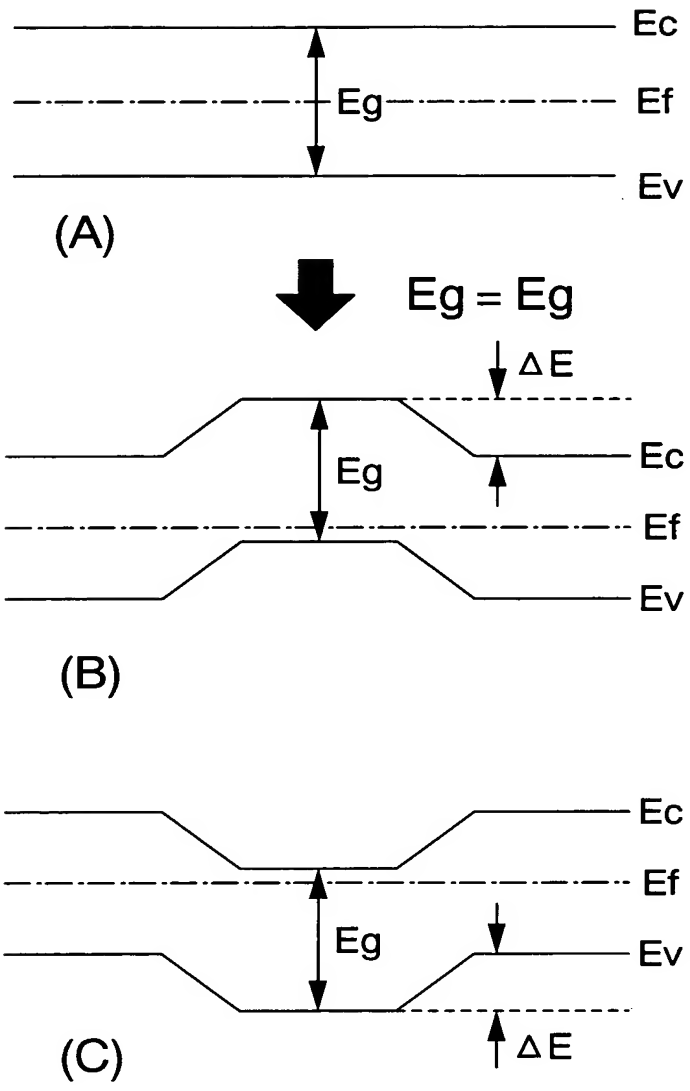
mode	Vd	Vs	V_{EG}	V_{CG}	mechanism
at writing	8	0	3	12	hot electron injection for floating gate
at erasing	-	0	20	0	F-N tunnel erasure from floating gate
at readout	1	0	0	5	-

[Reference No.] P003779-01
[Name of Document] Drawing
[Fig. 1]

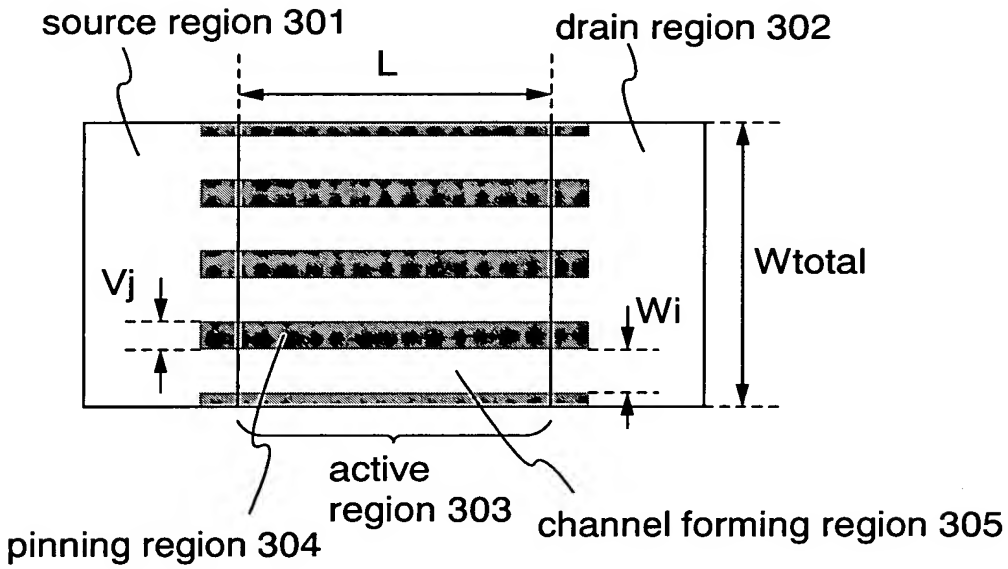


[Reference No.] P003779-01

[Fig. 2]

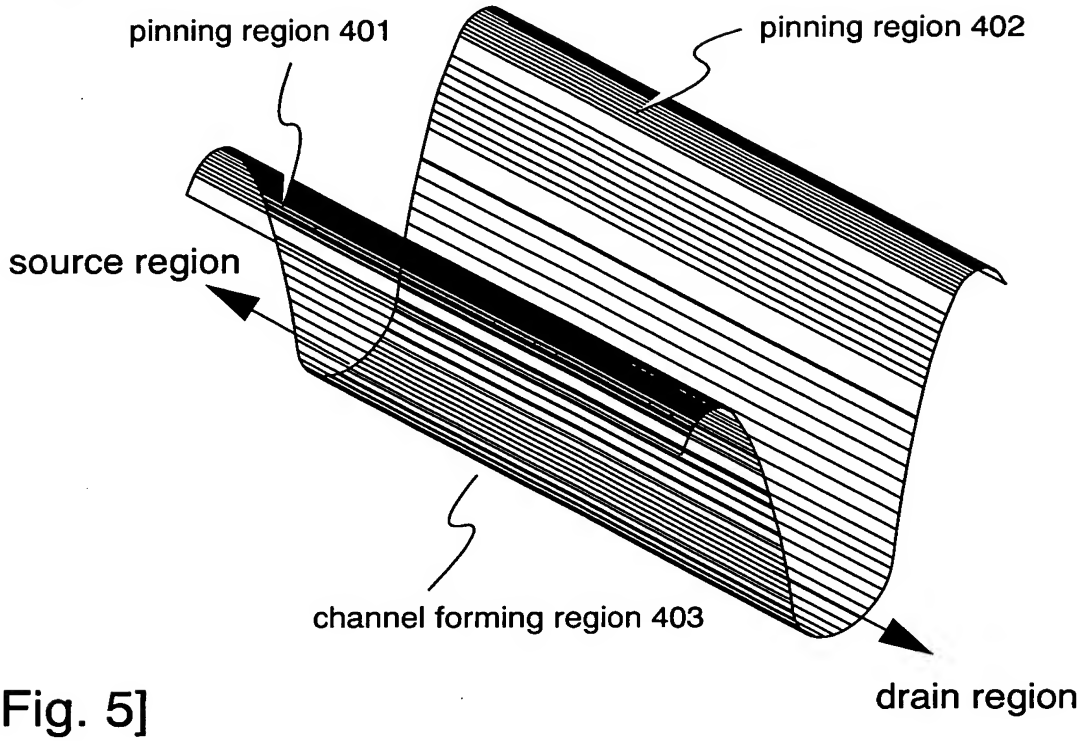


[Reference No.] P003779-01
[Fig. 3]

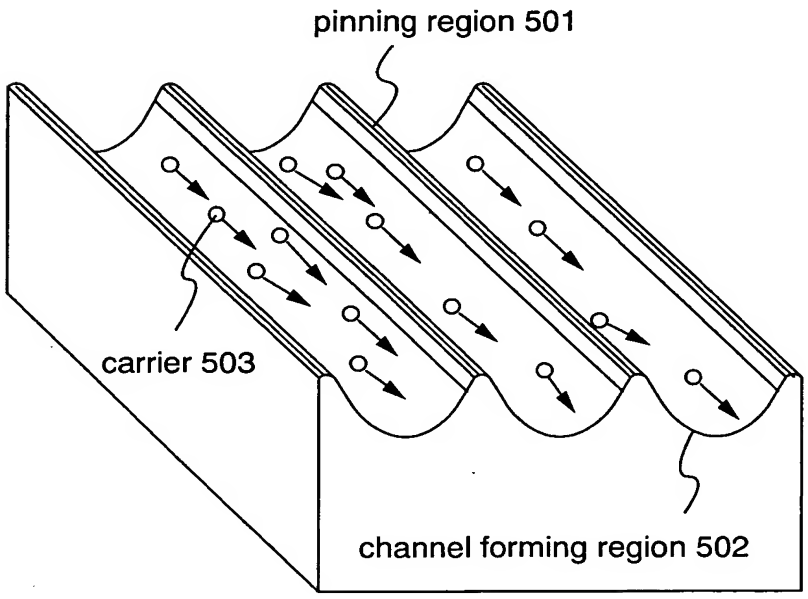


[Reference No.] P003779-01

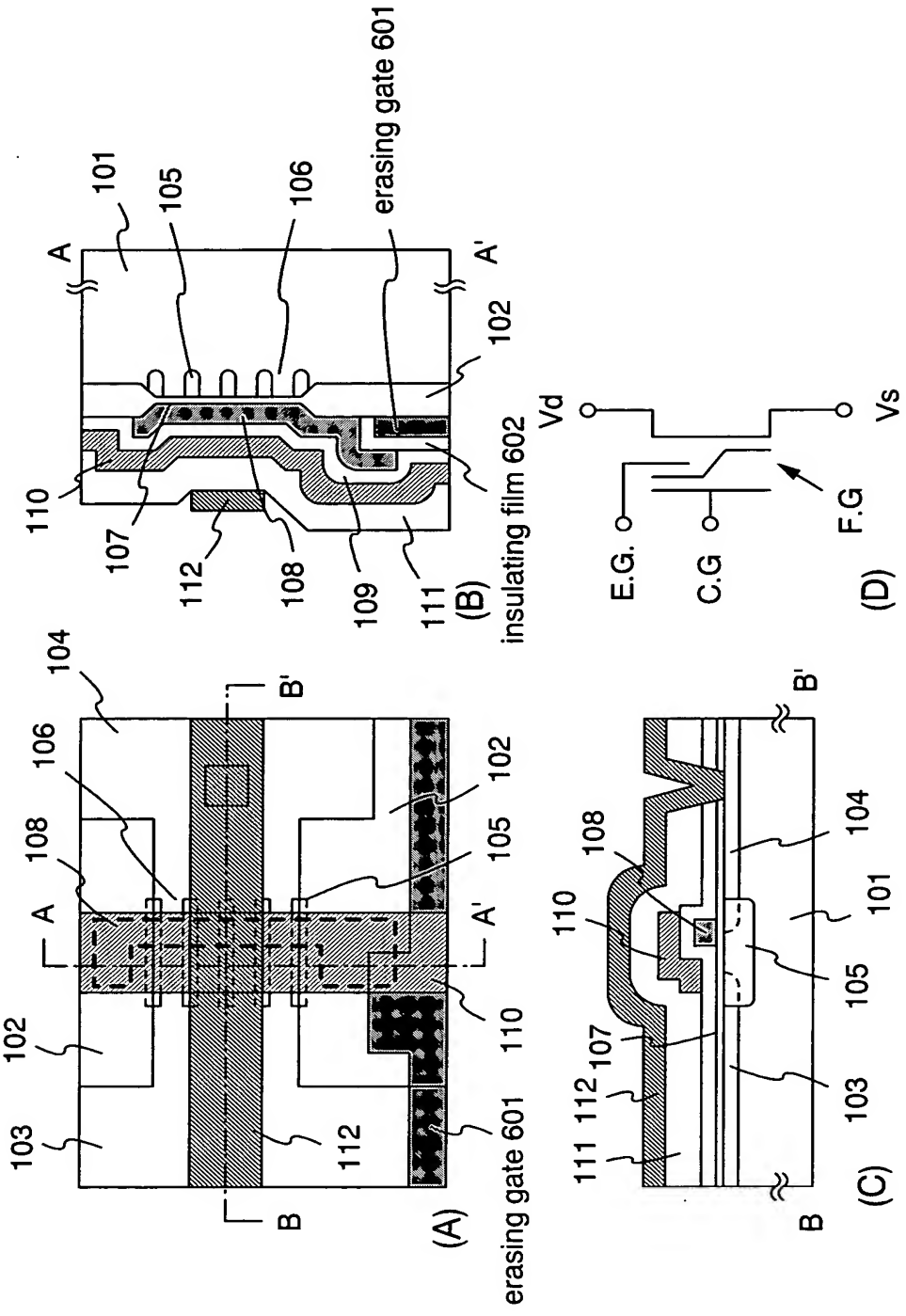
[Fig. 4]

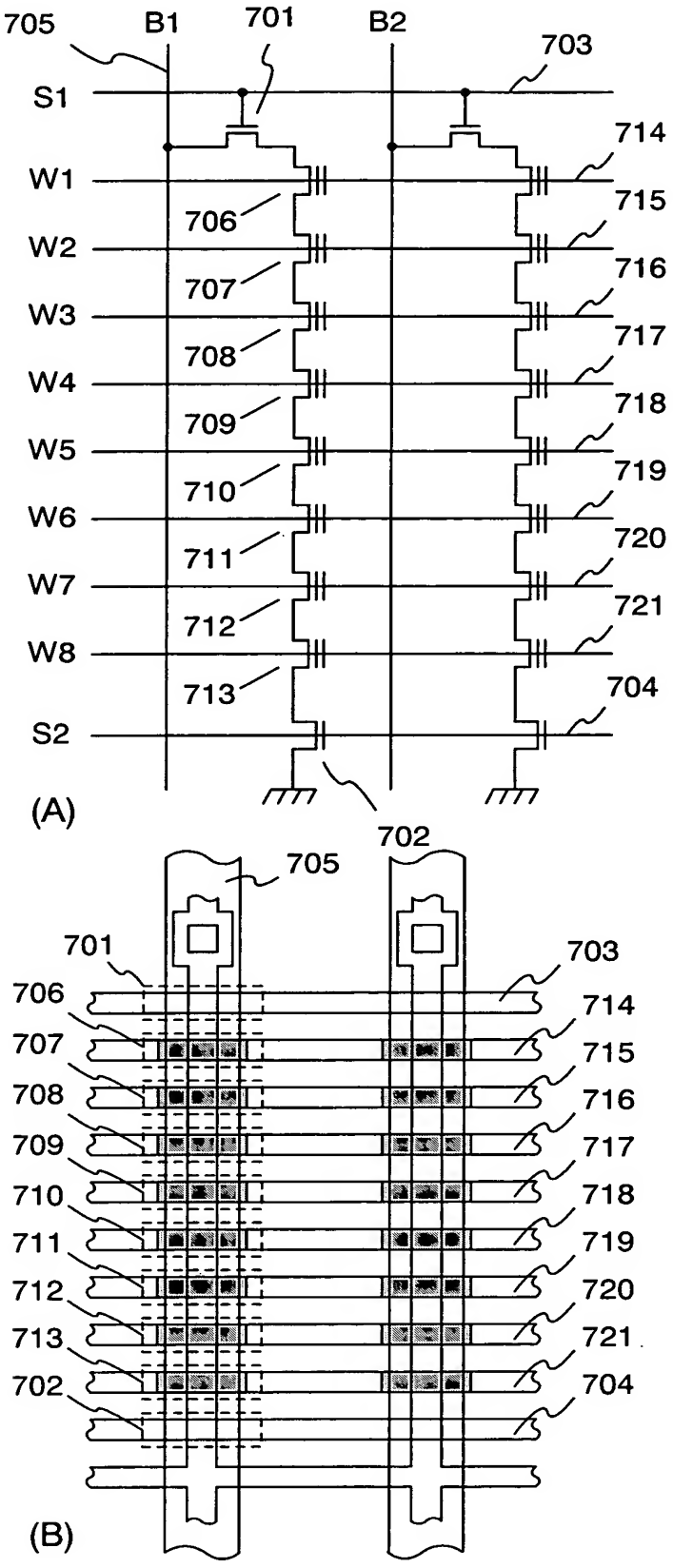


[Fig. 5]

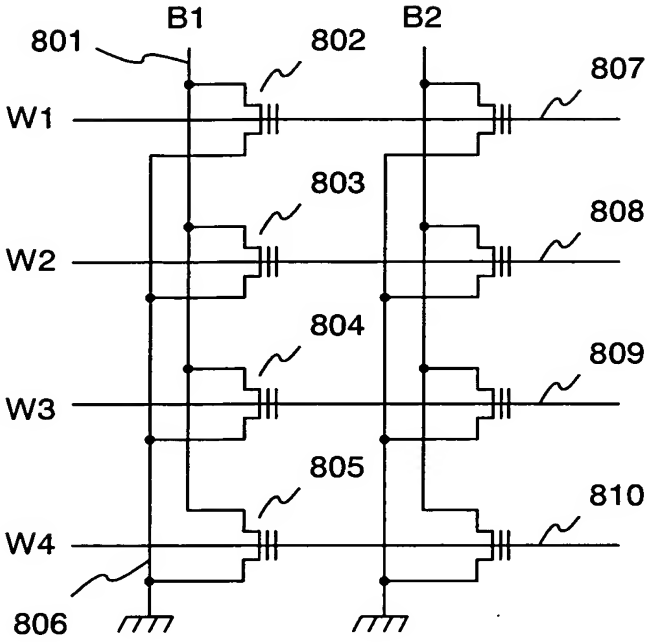


[Reference No.] P003779-01
[Fig. 6]

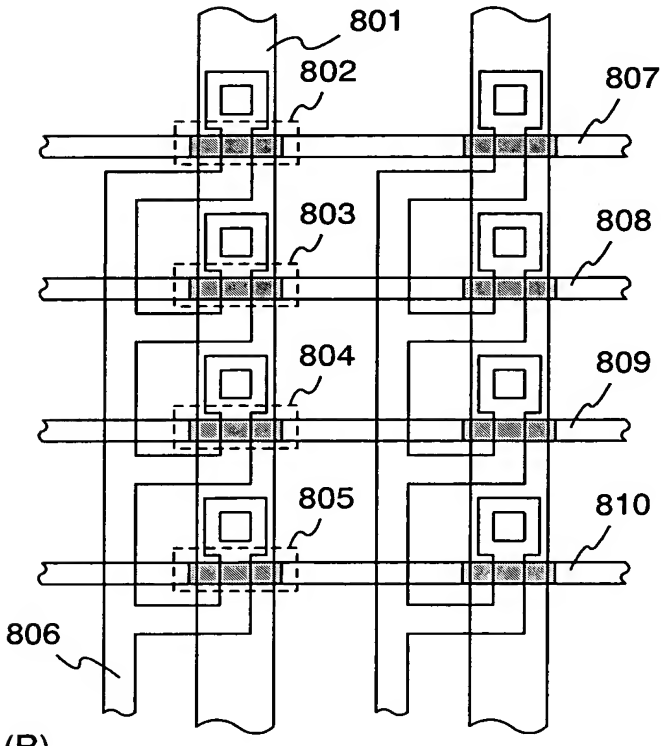



705
B1
701
B2
703
S1
714
W1
706
715
W2
707
716
W3
708
717
W4
709
718
W5
710
719
W6
711
720
W7
712
721
W8
713
704
S2
(A)
702
705
701
703
706
714
707
715
708
716
709
717
710
718
711
719
712
720
713
721
702
704
(B)

[Reference No.] P003779-01
[Fig. 8]



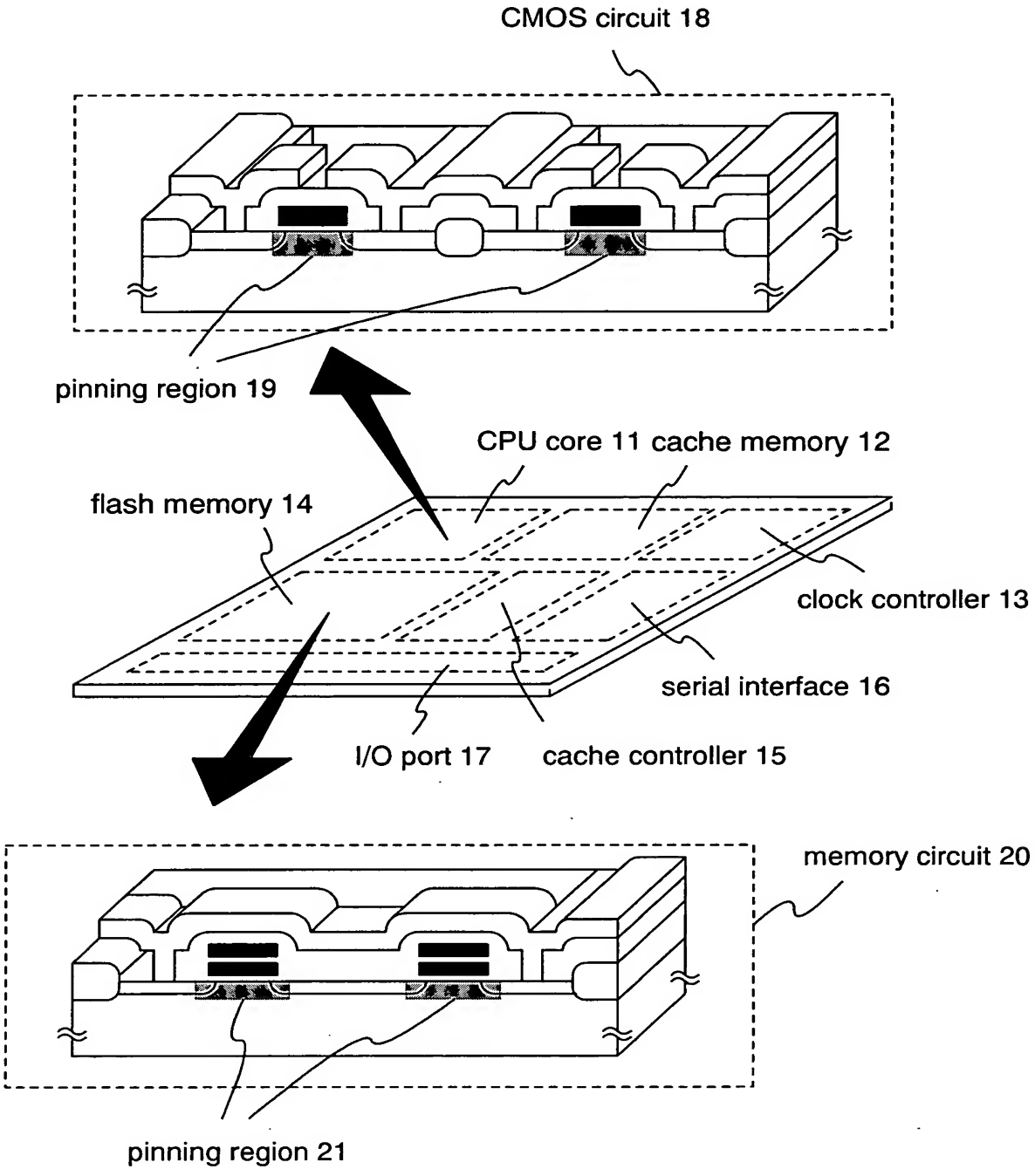
(A)



(B)

[Reference No.] P003779-01

[Fig. 9]



[Reference No.] P003779-01

[Fig. 10]